Frequency beamforming of dolphin whistles using a sparse three-element towed array

Aaron Thode[a)]
Marine Physical Laboratory, Scripps Institution of Oceanography, San Diego, California 92093-0205

Thomas Norris[b)] and Jay Barlow
Southwest Fisheries Science Center, National Oceanographic and Atmospheric Administration, P.O. Box 271, San Diego, California 92038

(Received 9 February 1999; accepted for publication 21 January 2000)

Acoustic bearings are obtained from dolphin whistles using frequency-domain (FD) beamforming techniques on signals recorded on a three-element 9-m aperture towed array. Due to the wide element separation, the high-frequency (kHz range) signals generate numerous grating lobes, but these lobes shift bearing with beamformed frequency, allowing identification of the true bearing whenever the whistles have over 1 kHz bandwidth. This method was validated by matching a sighting of a compact group of dolphins with acoustic bearing estimates. The system was subsequently used to detect and determine bearings from animals at least 3 km away and in Beaufort 5+ conditions. Frequency-domain beamforming has advantages over temporal cross correlation when the signals are faint and/or overlapping. © *2000 Acoustical Society of America.* [S0001-4966(00)00405-7]

PACS numbers: 43.80.Lb [WA]

INTRODUCTION

Acoustic bearings from kilohertz-range marine mammal vocalizations are typically obtained via two methods: either by time-delay (TD) measurements of signals recorded on widely spaced hydrophones,[1,2] or more recently[3] by beamforming in the frequency domain (FD), using signals collected from many closely spaced hydrophones. Although mathematically equivalent to delaying and summing signals in time, frequency-domain methods demonstrate two advantages over time-domain beamforming. First, FD methods enable greater coherent gain when used with multielement arrays[4] and thus should be able to obtain bearings from signals too weak to be localized by TD methods. Second, delphinid vocalizations often overlap in time and frequency. Under these circumstances time-domain methods may fail due to interference between the signals. Frequency-domain methods can isolate overlapping signals from each other, because they allow precise control over which frequency components in a data sample are processed.

A potential disadvantage of FD beamforming occurs whenever the spacing between adjacent hydrophones exceeds half an acoustic wavelength at a given frequency, and the beamformer cannot distinguish between the true signal bearing (the mainlobe) and multiple false bearings (i.e., grating lobes, or sidelobes). For example, FD beamforming on a 10-kHz pure tone (15-cm wavelength) with hydrophones spaced 5 m apart generates over 60 grating lobes, making mainlobe identification impossible.

Most dolphin whistles are frequency modulated,[5] and thus have wide frequency bandwidth. Because grating lobe maxima shift position with beamformed frequency, this bandwidth can be exploited to identify the mainlobe. In this letter these sidelobe properties were exploited to conduct FD beamforming on kilohertz-range marine mammal calls recorded on three hydrophones separated by 4 and 5 m, a situation that routinely generated 30–60 grating lobes at each frequency.

I. BACKGROUND

Conventional frequency-domain algorithms[4] decompose a received signal into its frequency components, and then generally assume that each component represents a plane wave arriving at a bearing θ from the array. In the following discussion a bearing of 0° refers to an angle arriving perpendicular (broadside) to the line array. The algorithm compares the data with that expected from a plane wave arriving from a test bearing θ', and produces the following output:

$$B(f,\theta)=\left|\sum_n e^{i(2\pi d/\lambda)(\sin\theta-\sin\theta')}\right|^2. \tag{1}$$

Here, d is the spacing between array hydrophones, and λ is the wavelength of the signal at frequency f. Note how the beamformer output will always attain a maximum (mainlobe) when $\theta=\theta'$. However, if the ratio d/λ is greater than 1/2, then other bearings θ'_m will produce beamformer maxima, or grating lobes. Their bearings are given by the following formula:

$$\sin\theta_m=\sin\theta\pm\frac{cm}{fd},\quad m=0,\pm 1,\pm 2,\pm 3,... \tag{2}$$

where c is the waterborne sound speed, f is the signal frequency, and m is an integer.

From Eq. (2) it follows that grating lobe bearings (which

[a)]Current address: Dept. of Ocean Eng., Room 5-212, MIT, 77 Massachusetts Ave., Cambridge, MA 02139.
[b)]Current address: SAIC Maritime Division, 3990 Old Town Ave., San Diego, CA 92110.

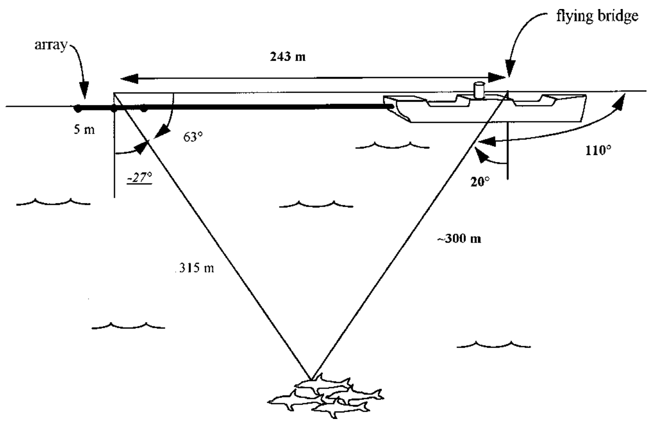

FIG. 1. Geometry used to convert a visual bearing into a relative array bearing for data in Fig. 2. Boldfaced numbers are coordinates measured by a visual observer on the flying bridge. The underlined number is the derived bearing, relative to the array.

have a nonzero value of m) will change with the beamformed frequency. Although the grating lobe magnitudes are identical to that of the mainlobe, the latter is easily identified as the maxima that maintains a fixed bearing with frequency. Averaging multiple beams generated at different frequencies thus suppresses the grating lobes. This fact allows FD methods to be applied to three widely spaced elements, preserving array aperture and thus angular precision.[4] Frequency averaging is a common technique in many applications involving sparse arrays.[6,7]

II. EXPERIMENTAL SETUP

The visual and acoustic data from dolphins were recorded in Nov./Dec. 1998 during a National Marine Fisheries Service dolphin survey cruise, near the coast of Panama.[8] A three-element array (built by Don Norris, SonaTech, Inc.) was deployed at ~5 m depth, 200 m behind the fantail of the R/V ENDEAVOR, a 175-ft oceanographic research vessel. The individual elements were spaced 4 and 5 m apart, and had omnidirectional sensitivity and a nearly linear frequency response from 2–150 kHz. The signals from the three phones were fed into a Mackie CR1604-VLZ audio mixer, which was used as a 200-Hz high-pass filter. The signals were then passed through a TASCAM DA-38 digital recorder with a 48-kHz low-pass filter, and recorded for later analysis. Simultaneously, the TASCAM outputs were passed through the mixer a second time to allow further amplification and filtering (usually to emphasize the 8–20 kHz band), before being sampled into two Data Translation DT-3809 12-bit A/D cards within a Dell Optiplex 200-MHz Pentium computer.

The beamforming software was written in MATLAB 5.0. When executed, the program sampled the data at 48 kHz and wrote up to a minute of multichannel data to hard disk. The data from the first hydrophone were then read and displayed as a spectrogram. A simple mouse-driven graphical user interface allowed the user to trace the time-frequency contour of a whistle in order to analyze it. When selected, the program computed and averaged the spectra of the first 1024 points of the signal from each hydrophone. The frequency with the highest average power within a certain frequency window was then selected for beamforming. The program then adjusted its selection window based on the contour trace, obtained the next data sample, and repeated the process. Using this technique, most of the frequency components of a modulated sweep could be beamformed, regardless of whistle contour or duration. A selection window of ± 500 Hz bandwidth was found sufficient to track all but the most rapidly modulated whistles. In cases where multiple whistles were recorded at the same time, the selection window was made narrower to avoid interference.

The individual beamformer outputs were then displayed as a frequency-azimuth plot before being incoherently averaged to yield a total beamformer output, where the mainlobe is revealed as the global maximum.

III. RESULTS

The beamforming algorithm was first calibrated on sperm whale clicks arriving from known bearings, identified using software developed by J. Barlow. Next, the methods were tested on over 100 whistles and burst pulses. On 16:36 Dec. 6, the system determined bearings from animals that were 3 km away, as determined by visual sighting methods. On another occasion the acoustic observers obtained bearings under Beaufort 5+ conditions, and assisted visual observers in locating a school of *Tursiops truncatus* during twilight, providing an example of how acoustic bearings can help supplement visual methods under adverse conditions.

Three specific examples of the system output are now presented. The first example demonstrated a match between a visual and acoustic observation on Dec. 5, 1998. On this date a compact group of 36–45 striped dolphins (*Stenella coeruleoalba*) approached the ship, rode the bow, and eventually passed by the beam of the Endeavor. The occurrence of an isolated, compact group of animals was an uncommon event—typically, dolphin encounters consisted of hundreds of animals widely scattered over an area. Between 10:00 and 10:08 an experienced marine mammal observer tracked the group from the flying bridge using 25x reticulated binoculars, recording the horizontal and vertical azimuth from the horizon approximately every 30 sec. At 10:02 the animals were at a horizontal azimuth of 20° past the ship's beam (110° relative to the ship's bow), and about 300 m away from the flying bridge. The bearing of the group relative to the array was estimated to be near $-26°$ (as illustrated in Fig. 1).

A frequency-modulated (FM) whistle was recorded at 10:02, and is displayed in Fig. 2(a). Over 1 sec the whistle swept between a frequency of 9.8 to 18 kHz. The individual beams obtained from this whistle are plotted vs frequency in Fig. 2(b). The rapid frequency modulation precludes beamforming at every frequency bin within this band; however, the identity of the mainlobe is clear. Note that even if one were restricted to beamforming between 9–11 kHz, the mainlobe location would be readily distinguishable. In fact, signals with as little as 1 kHz modulation were successfully beamformed.

By averaging all the beams together [i.e., summing along the frequency axis in Fig. 2(b)], and subtracting the

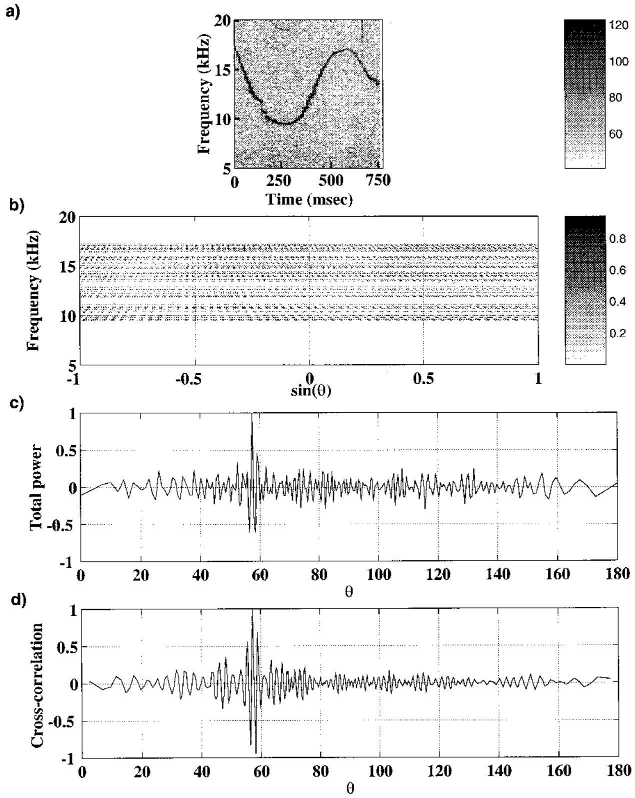

FIG. 2. (a) Spectrogram of dolphin whistle recorded at 10:02 am, Dec. 5, 1998. Fourier transform length is 1024 pts, sampling rate is 48 kHz, window overlap is 50%. (b) Plane-wave (PW) beams from (a), plotted as a function of frequency and the sine of the bearing. A negative value means the signal is arriving from the ship's bow. Dark colors represent stronger correlation. Note how all maxima shift positions with frequency, except for the mainlobe at position $\sin(\theta)=-0.52$. (c) Incoherent average of all beams vs bearing θ. The whistle is arriving from a bearing of $-33°$, or 57° away from the ship's heading. The mean value of the averaged beams has been subtracted and the pattern normalized by its maximum. (d) Cross correlation of the entire time series in (a), prefiltered between 9.4 and 17 kHz.

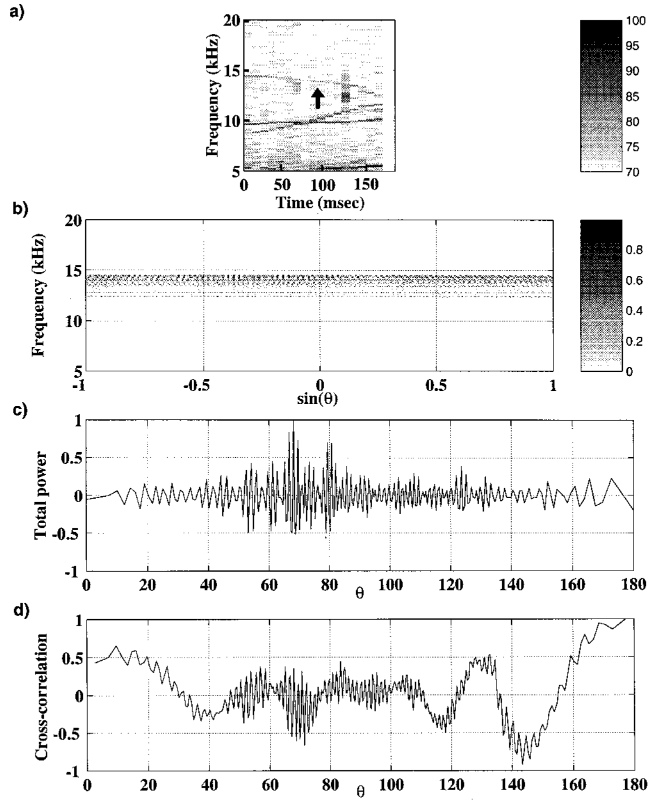

FIG. 3. (a) Spectrogram of dolphin whistle at 17:16, Dec. 6, 1998. Fast Fourier transform (FFT) length is 1024 pts, sampling rate is 48 kHz, window overlap is 50%. The arrow indicates the faint whistle to be processed. (b) PW beams plotted as a function of frequency and $\sin(\theta)$. (c) Incoherent summation of the beams displayed in (b), vs bearing θ. (d) Cross correlation of the entire time series in (a), prefiltered between 12 and 15 kHz.

mean, a combined beamformer output can be displayed [Fig. 2(c)]. (Note that the horizontal axis is now in terms of θ, not $\sin\theta$.) Finally, the result of cross correlating the first two hydrophone signals is shown in Fig. 2(d). For isolated whistles with high signal-to-noise ratio, both the time- and frequency-domain methods give similar results.

The beamformer estimates an acoustic bearing of $-33°$, whereas the visual observer bearing estimate was $-26°$. Given the disparity in timing between the visual and acoustic records (~10 sec), and the rapid motion of the dolphin group, the two measurements are in close agreement. A second visual observation at 10:08 gave a bearing of 50°, and the corresponding acoustic bearing estimate was 41°—again, an approximate match. Despite the fact that the range of the animals was less than 400 m, the plane-wave approximation seemed to work well, and modeling wave front curvature was not required.

The next example shows the advantage of using coherent gain from three hydrophones. Figure 3 illustrates a faint whistle recorded at 17:16 on Dec. 6, one day after the previous example. At this time multiple animals in a group were whistling; this example concerns the whistle indicated by the arrow in Fig. 3(a). As in Fig. 2, the individual beamformed outputs are shown in Fig. 3(b) and the averaged beamformer output between 12 and 15 kHz is plotted in Fig. 3(c) For comparison, the cross correlation between the first two phones is provided in Fig. 3(d). In this last plot the data have been digitally bandpass filtered using a finite impulse response (FIR) filter, designed to bandpass the same frequency range displayed in Fig. 3(b). The FD result yields a bearing of 68°, whereas cross correlation fails due to the low signal-to-noise ratio. If the original signal is digitally refiltered using different choices for passband, the cross correlation still fails.

The final example was recorded at 17:02 Dec. 6, and is shown in Fig. 4. It illustrates a common situation, wherein two whistles by two separate animals overlap in frequency and time, such that the signals cannot be isolated by simple bandpass filtering. One whistle has the upward U-shaped contour labeled ''1'' in Fig. 4(a) and the second has the relatively longer FM sweep labeled ''2'' in the same spectrogram. When applying the beamformer algorithm, a frequency selection window of 100 Hz was used to ensure that the beamformer sampled only the energy from whistle 2, whose output is shown in Fig. 4(b) and (c). The mainlobe position indicates that whistle 2 is arriving from the bow of the ship.

When time-domain techniques are applied to two overlapping signals, one obtains two separate peaks if the signals are uncorrelated, with no means of telling which peak was generated by which signal. If the signals are correlated (i.e.,

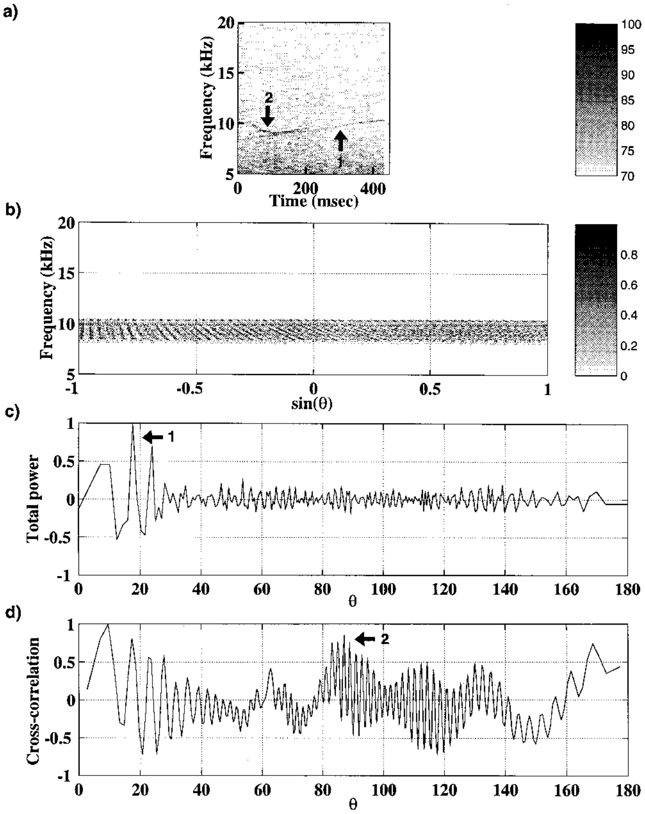

FIG. 4. (a) Spectrogram of overlapping dolphin whistles at 17:02, Dec. 6, 1998. For this case the search window is narrowed to 0.1 kHz around the frequency contour of whistle #1. (b) PW beams from whistle 1, plotted as a function of frequency and $\sin(\theta)$. (c) Incoherent summation of the beams displayed in (b), vs bearing θ. The arrow indicates the bearing of the whistle 1. (d) Cross correlation of the time series in (a), prefiltered between 8 and 11 kHz. As it is impossible to filter out the effects of whistle 2, which is arriving from broadside, the output is corrupted and the bearing of whistle 1 is not obtained. The arrow indicates the contamination from whistle 2.

both share similar frequency contours) then the bearing estimates can be compromised, as in the case in Fig. 4(d). Here the time-domain output shows the combined effects of whistle 2, which is arriving from broadside, and whistle 1. The bearing estimate of the latter whistle has been corrupted due to the fact that both signals are slightly correlated. The FD beamformer avoided this problem and yielded separate bearings for each whistle. The ability to obtain bearings from overlapping whistles was useful when trying to separate distant animals ahead of the ship from nearby animals the ship had already passed.

IV. CONCLUSION

Successful frequency-domain beamforming on dolphin whistles has been demonstrated on a three-element sparse array, in which the array elements are spaced apart by more than 50 times a typical acoustic wavelength. FD beamforming is demonstrated to be more effective than temporal cross correlation in cases where the signal is faint and/or overlapping with other signals. The sparse-array beamforming methods discussed here and in Ref. 3 shouldwork on most broadband marine mammal sounds, including the majority of tonal sounds produced by odontecetes.[9]

ACKNOWLEDGMENTS

The authors would like to thank the crew and the marine mammal observers of the R/V ENDEAVOR (particularly Laura Morse, Meghan Donahue, and Jorge Rodrîquez) for their cooperation and support during the field testing. They would also like to thank Patrick Miller and Peter Tyack of Woods Hole Oceanographic Institute for advice about the manuscript. This work was funded under U.S. Department of Commerce Requisition #NFFR2100800295. Equipment support was provided by the National Marine Fisheries Service and the Office of Naval Research. Tim Gerrodette was Chief Scientist for the dolphin survey.

[1] L. E. Freitag and P. L. Tyack, ''Passive acoustic localization of the Atlantic bottlenose dolphin using whistles and echolocation clicks,'' J. Acoust. Soc. Am. **93**, 2197–2205 (1993).

[2] W. A. Watkins and W. E. Schevill, ''Sound source location by arrival times on a non-rigid three-dimensional hydrophone array,'' Deep-Sea Res. **19**, 691–706 (1972).

[3] P. J. Miller and P. L. Tyack, ''A small towed beamforming array to identify vocalizing resident killer whales (Orcinus Orca) concurrent with focal behavioral observations,'' Deep-Sea Research II **45**, 1389–1405 (1998).

[4] D. J. DeFatta, J. G. Lucas, and W. S. Hodgkiss, ''Conventional beamforming, Appendix A11,'' in *Digital Signal Processing: A System Design Approach* (Wiley, New York, 1988), pp. 629–649.

[5] J. N. Oswald, ''Delphinid whistles recorded in the eastern tropical Pacific Ocean,'' National Marine Fisheries Service, Southwest Fisheries Science Center, Administrative LJ-99-07C, July 1999.

[6] M. J. Hinich, ''Processing spatially aliased arrays,'' J. Acoust. Soc. Am. **64**, 792–794 (1978).

[7] F. Anderson, W. Christensen, L. Fullerton, and B. Kortegaard, ''Ultra-wideband beamforming in sparse arrays,'' IEE Proc., Part H: Microwaves, Antennas, Propag. **138**, 342–346 (1991).

[8] P. Olson and T. Gerrodette, ''Report of the meeting to review the preliminary estimates of Eastern Tropical Pacific dolphin abundance in 1998,'' National Marine Fisheries Service, Southwest Fisheries Science Center, Administrative LJ-99-03, Jan. 21, 1999.

[9] J. N. Mathews, L. E. Rendall, J. C. D. Gordon, and D. W. Macdonald, ''A review of frequency and time parameters of cetacean tonal calls,'' Bioacoustics **10**, 47–71 (1999).